FOUNDATIONS
ADVANCES

Volume 77 (2021)

Supporting information for article:

New perspectives in macromolecular powder diffraction using single photon counting strip detectors: High resolution structure of the pharmaceutical peptide, Octreotide

Maria Spiliopoulou, Fotini Karavassili, Dimitris-Panagiotis Triandafillidis, Alexandros Valmas, Christos Kosinas, Kleomenis Barlos, Kostas K. Barlos, Mickael Morin, Mathilde L. Reinle-Schmitt, Fabia Gozzo and Irene Margiolaki

S1. Supplementary Material

An exhaustive modelling of the beamline optics, diffractometer and detectors has already been performed in 2006 (Gozzo *et al.,* 2006 and references herein) and 2010 (Gozzo *et al.,* 2010 and references herein). In spite of the 2011 optics upgrade undergone by the Materials Science beamline, these modelling studies are still relevant references as the optics always consists of 2 mirrors and 2 crystals. In line with the 2006 and 2010 studies, a set of Instrumental Resolution Functions (IRF) curves, i.e. FWHM versus 2-theta curves has been constructed, which describes the effect on IRFs of a highly-focused beam and the sub-strip data rebinning.

IRF curves are ideally constructed using reference standards for peak profile analysis with diffraction peaks well distributed in the 2-theta range of interest, whose intrinsic contribution to the overall peak profile is negligible, so to reflect the pure instrumental contribution to the peak profile. In 2006 and 2010, the $Na_2Ca_3Al_2F_{14}$ (NAC) and Si NIST640C reference powders were used, respectively. NAC is not a certified reference standard, but widely used in the high-resolution synchrotron powder diffraction community for its negligible contribution to the peak profile. With multicrystal analyzer detector data, IRF curves with NAC were modelled with 2 μm crystallite size contribution (Gozzo *et al.*, 2006) and are expected to generate a negligible intrinsic contribution to the peak profile with position sensitive detectors, like MYTHEN II detector, due to the dominant effect of the sample dimension to the overall peak profile in such case. Absorption effects have, however, been observed that generate apparent peak profile broadening effects, especially at lower photon energies and low 2-theta angles, which are a function of the capillary diameter.

In the case of applications to proteins, characterized by very large unit cell volumes, the 2-theta angular range of interest is, however, the one corresponding to d-spacing above 3.5-4 Å. This makes the choice of NAC or Silicon not ideal as lineshape profile reference materials as both have a very limited number of peaks above d= 3.5-4 Å (only 1 peak for silicon, 4-5 peaks for NAC, no peak above 7 Å).

In order to generate meaningful IRF curves for applications to proteins, we have therefore performed parametrical single peak fits using diffraction patterns recorded on two highly crystalline protein powders, measured during 2 distinct experimental sessions with nominally the same experimental set up (highly-focused beam) and acquisition protocol (sub-strip rebinning) described in the article. The first crystalline protein is the same Octreotide peptide protein powder discussed in this article (OCT_S1, λ= 1.3004392(8) Å) and the second is the crystalline powder of human insulin (NL4_1, λ =1.3831050(6) Å). In order to construct meaningful IRF curves, only peaks that are single reflections (as ascertained by Pawley refinements) with a symmetric peak profile or a peak profile straightforwardly modelled by Finger–Cox–Jephcoat (FCJ) function (Finger et al., 1994) have been retained. This selection condition does not guarantee that the reflections retained for the IRF curves uniquely reflect the instrumental

contribution to the peak profile. However, a contribution to the peak profile by intrinsic broadening would sum up to the overall FWHM of our IRF curves. Therefore, the resolution function described by the IRF curves so-constructed describes the "worst" angular resolution as a function of the 2-theta angles for the working photon energy of approximately 9 keV that can be achieved with our experimental set up and data acquisition strategy.

Figure S1 shows the Instrumental Resolution Function curves, i.e. FWHM values as a function of 2-theta values for the OCT_S1 and NL4_1 proteins patterns loaded in 0.8 mm glass capillaries. The FWHM values as a function of 2-theta for NAC_0.3 mm and NAC_0.8 mm have been added to the plot as reference values. We observe the effect of apparent peak broadening of NAC that we attribute to absorption effects, especially significant in the 0.8 mm NAC IRF curve. In the region between 10 and 20 degrees in 2-theta (light yellow box) the FWHM values of peaks of the 2 crystalline proteins are substantially smaller than the values refined for NAC in 0.8 mm capillary diameter and consistent with the FWHM values refined for NAC 0.3 mm capillary diameter. We interpret this observation in terms of apparent peak broadening due to absorption affecting NAC peaks, suggesting NAC should be substantially diluted with a low-Z powder if to be used as reference lineshape standard in capillaries larger than 0.3mm diameter.

Figure S1 Instrumental Resolution Function (IRF) curves constructed to describe the resolution conditions under which data on proteins were recorded. The IRF below 15 degrees in 2-theta at the working photon energy of approximately 9 keV (d-spacing > 4.5-5 Å) is described by the FWHM values as obtained with single peak line shape parametrical fits performed on highly crystalline protein powders. Only single reflections with a symmetrical profile and/or single reflections with axial divergence asymmetry well modelled by Finger–Cox–Jephcoat (FCJ) asymmetric function were retained to construct the IRF curves, as described in the text.

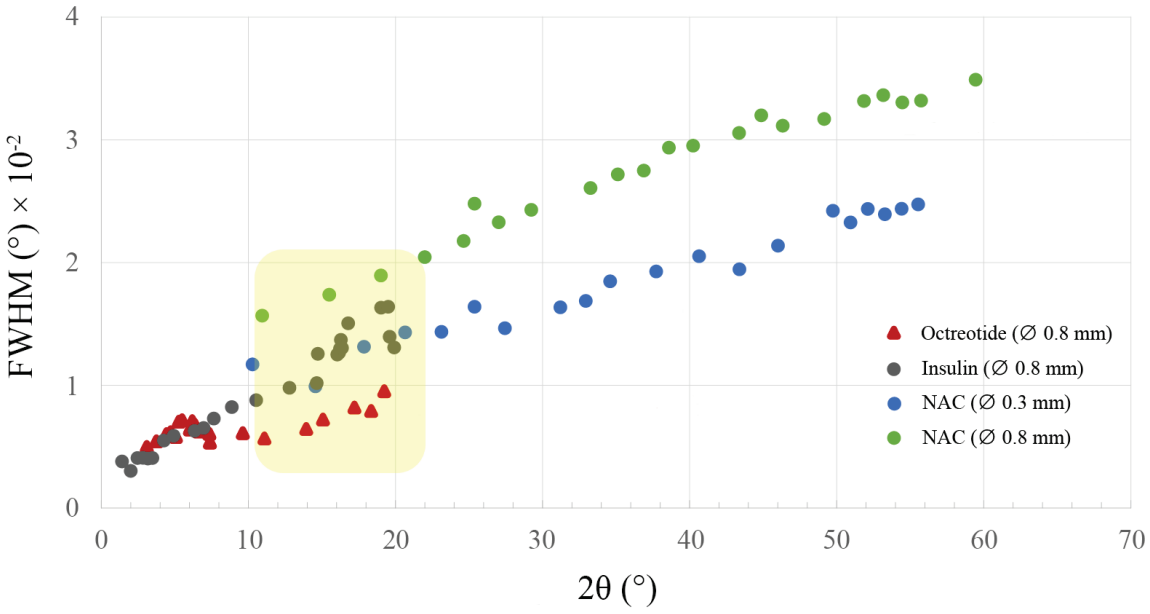

Figure S2 Secondary structure of chains in the asymmetric unit (top left). Chain *A* and part of chain *B* form an anti-parallel *β*-sheet. By overlaying the three different chains, their structural differences are highlighted (middle and bottom left). Detailed secondary structure is reported in Table 3. Intramolecular interactions of chains in the same orientation are shown on the right. Numbers correspond to IDs reported in Table 2.

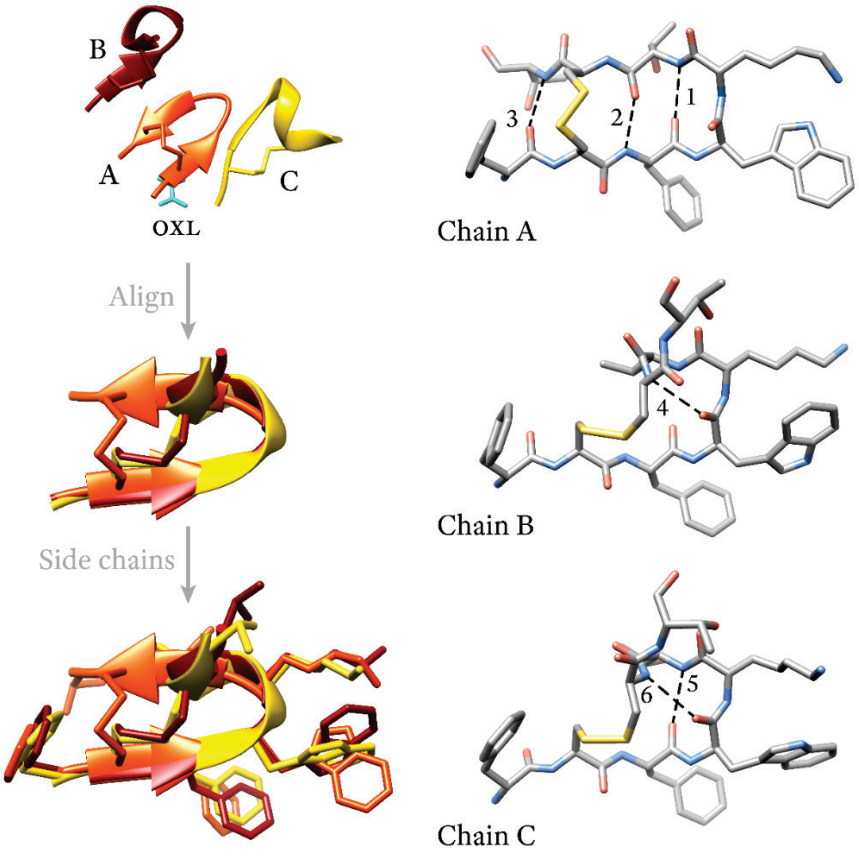

Figure S3 Selected amino acids of the refined structures (upper: Fili et al., 2019; lower: 6VC1), illustrating a comparison in model fitting between the two structures with the total *OMIT* maps at 1σ. The individual maps were computed at the final stage of the analysis and was generated using *SFCHECK* software.